Влияние качества библиографического описания на библиометрические оценки¹

В. А. Цветкова[✉], Ю. В. Мохначева

**Цветкова
Валентина
Алексеевна,**

Библиотека
по естественным наукам
Российской академии
наук, ул. Знаменка, 11/11,
Москва, 119991, Россия,

доктор технических наук, профессор,
главный научный сотрудник БЕН РАН

ORCID: [0000-0002-2796-3574](https://orcid.org/0000-0002-2796-3574)
e-mail: Vats08@mail.ru

**Мохначева
Юлия Валерьевна,**

Библиотека по естествен-
ным наукам Российской
академии наук,
ул. Знаменка, 11/11,
Москва, 119991, Россия,

кандидат педагогических наук,
ведущий научный сотрудник БЕН РАН

ORCID: [0000-0001-5780-485X](https://orcid.org/0000-0001-5780-485X)
e-mail: j-v-m@yandex.ru

Аннотация. Рассмотрено влияние качества библиографического описания на результаты поиска в библиометрических базах данных. Предложены три базовых уровня для библиометрических оценок. Показано влияние библиометрических погрешностей на качество оценок на каждом уровне. Отмечено, что наиболее чувствителен к погрешностям уровень, на котором работа проводится с конкретными авторами. Ответственность за некорректность библиографических данных в наукометрических поисковых системах лежит на генераторах информационных ресурсов этих систем.

Ключевые слова: библиометрические оценки, библиографические описания, публикации, качество поиска

Для цитирования: Цветкова В. А., Мохначева Ю. В. Влияние качества библиографического описания на библиометрические оценки // Библиосфера. 2021. № 2. С. 59–64. <https://doi.org/10.20913/1815-3186-2021-2-59-64>.

Статья поступила в редакцию 17.05.2021
Получена после доработки 20.05.2021
Принята для публикации 24.05.2021

© В. А. Цветкова, Ю. В. Мохначева, 2021

¹ Научная статья подготовлена на основе доклада на III Международном библиографическом конгрессе, 27–30 апреля 2021 г.

Influence of the quality of the bibliographic description on bibliometric scores

Valentina A. Tsvetkova✉, Yuliya V. Mokhnacheva

Tsvetkova Valentina Alekseevna,

Library for Natural Sciences
of the Russian Academy of
Sciences,
Znamenka str., 11/11, Moscow,
119991, Russia,
Doctor of Sciences (Techn.),
Professor, Chief Researcher

ORCID: [0000-0002-2796-3574](https://orcid.org/0000-0002-2796-3574)

e-mail: Vats08@mail.ru

Mokhnacheva Yuliya Valerjevna,

Library for Natural Sciences
of the Russian Academy of
Sciences,
Znamenka str., 11/11, Moscow,
119991, Russia,
Candidate of Pedagogic Sciences,
Leading Researcher

ORCID: [0000-0001-5780-485X](https://orcid.org/0000-0001-5780-485X)

e-mail: j-v-m@yandex.ru

Abstract. The influence of the quality of the bibliographic description on the search results in bibliometric data bases is considered. Three basic levels for bibliometric assessments are proposed. The influence of bibliometric errors on the quality of assessments at each level is shown. It is noted that the level at which the work is carried out with specific authors is the most sensitive. The responsibility for the inaccuracy of bibliographic data in scientometric search engines lies on the generators of information resources of these systems.

Keywords: bibliometric assessments, bibliographic descriptions, publications, search quality

Citation: Tsvetkova V. A., Mokhnacheva Yu. V. Influence of the quality of the bibliographic description on bibliometric scores. *Bibliosphere*. 2021. № 2. P. 59–64. <https://doi.org/10.20913/1815-3186-2021-2-59-64>.

Received 17.05.2021

Revised 20.05.2021

Accepted 24.05.2021

Введение

В 1939 г. Джон Десмонд Бернал ([Bernal, 1939](#)) сформулировал основные характеристики, присущие «науке о науке», которую сегодня мы называем науковедением: внимание к проблемам организации науки и научно-исследовательского труда; широкое использование статистических данных о научных кадрах, учреждениях и публикациях (в настоящее время это библиометрические методы); анализ материалов о финансовом обеспечении науки и эффективности научных исследований; стремление постоянно оценивать и анализировать социальные функции науки, а также взаимосвязь и взаимообусловленность процессов ее развития с другими сторонами жизни общества ([Цветкова и др., 2018](#)).

Оценка научной деятельности с использованием библиометрического инструментария пронизывает все слои научного сообщества, включая непосредственно ученых и управленцев всех уровней, независимо от их понимания существа вопроса. В профессиональной печати опубликовано немало статей, посвященных анализу публикационной активности на основе библиометрических оценок ([Арчаков и др., 2013](#); [Березкина, Хренова, 2010](#); [Гиляревский, 2014](#); [Гуськов, 2015](#); [Лаврик, Глухов, 2015](#); [Маркусова, 2014](#); [Маршак-Шайкевич, 2013](#); [Москалева, 2013](#); [Мохначева, Харыбина, 2010](#)). В своих работах ([Мохначева,](#)

[Цветкова, 2017](#); [Цветкова, 2015](#)) мы акцентировали внимание на вопросах, связанных с качеством получаемых библиометрических оценок.

Основными инструментами для проведения библиометрических исследований в России служат три политематических ресурса: Web of Science Core Collection (WoS CC), Scopus и Российский индекс научного цитирования (РИНЦ). Эти ресурсы помимо информационно-библиографической информации предоставляют библиометрические данные, которые требуют особой осторожности при использовании. Кроме отсутствия у пользователя необходимых компетенций при поиске библиометрической и наукометрической информации, основная опасность кроется в ошибках, допущенных в сведениях о публикациях в базах данных ([Гуреев и др., 2017](#); [Мохначева, Цветкова, 2017](#)).

1. Спорные моменты инструментов для проведения библиометрических исследований в Web of Science Core Collection (WoS CC), Scopus и Российском индексе научного цитирования (РИНЦ)

Остановимся на некоторых спорных моментах этих инструментов, с которыми мы сталкиваемся при проведении поисков и анализе результатов с попыткой понять, насколько мы можем доверять данным этих систем и обходиться без экспертных

оценок. В этом исследовании мы не ставим цель критиковать аналитические системы: все изложенное характерно для всех информационных баз данных (систем), но для цитатных – библиометрических – становится критичным. Библиометрические данные напрямую влияют на научную политику в стране, от них зависит финансирование исследований, а также научная карьера ученых. Поэтому особенно важно максимально уменьшить количество ошибок и неточностей в базах данных (БД).

Существует три базовых уровня, по которым ведутся библиометрические исследования:

Макроуровень. Подразумевает библиометрический анализ публикационных массивов на уровне государств, в которых ведутся активные научные исследования. На макроуровне ввиду больших данных погрешности не принципиальны и нивелируются за счет значительной величины массива получаемой информации. В качестве примера приведем данные по динамике российского публикационного массива за 2010–2019 гг. БД WoS CC и Scopus (рис. 1).

По данным WoS CC, Россия с 2010 по 2020 г. по числу публикаций поднялась с 18-й позиции на 12-ю; по данным Scopus – с 16-го на 9-е место. Кроме того, и в РИНЦ наблюдается стабильный рост публикаций: в 2020 г., по сравнению с 2019 г., он составил 9 %.

Мезоуровень. На этом уровне проводится библиометрический анализ публикационной активности отдельных научных организаций (коллективов). Из-за того, что эти массивы уже намного меньше, чем по стране в целом, здесь мы можем столкнуться уже со значительными погрешностями. Неточности вызваны прежде всего вари-

антами названия организаций, вариантами перевода на английский язык, переименованием организаций, изменением ведомственной принадлежности, слиянием нескольких организаций и т. д.

Пример результата библиометрического анализа на мезоуровне приведен в [таблице](#), где представлены российские организации – лидеры в области физико-химической биологии.

Эти данные показывают, что Московский государственный университет имени М. В. Ломоносова представлен как в целом (все факультеты и НИИ), так и в отдельности – по факультетам. Это подтверждает наш тезис о том, что на мезоуровне важно правильное написание организаций.

Библиометрический анализ публикационных массивов научных организаций позволяет оценить не только вклад учреждения в публикационный массив страны, но и уровень влияния публикаций организации по цитируемости. При этом необходимо учитывать кадровый состав научных организаций, так как число научных сотрудников напрямую влияет на результаты и окончательную оценку.

Микроуровень. На этом уровне производится библиометрический анализ публикационной активности отдельных ученых. Именно здесь возникают максимальные погрешности, которые вызваны различными причинами.

Остановимся на некоторых из них.

В первую очередь это качество библиографических записей. Принцип GIGO (англ. Garbage In, Garbage Out, то есть «мусор на входе – мусор на выходе») особенно критичен. Он означает, что в системах, претендующих на получение статистических данных

Рис. 1. Динамика доли публикаций с российским участием в мировом массиве по БД WoS CC и Scopus за 2010–2019 гг. Все типы публикаций

Fig. 1. Dynamics of publication share with Russian participation in world publication array according to DB WoS CC and Scopus during 2010–2019.

All types of publications

Таблица. Российские организации – лидеры в области физико-химической биологии за 2010–2020 гг. по БД WoS CC (все типы публикаций)

Table. Russian leaders (organizations) on physical-chemical biology in 2010–2020 according to WoS CC

Научно-исследовательские учреждения	Количество публикаций	Средняя цитируемость одной публикации
Московский государственный университет имени М. В. Ломоносова (все факультеты и научно-исследовательские институты (НИИ))	6299	14,56
Институт биоорганической химии им. академиков М. М. Шемякина и Ю. А. Овчинникова Российской академии наук (РАН)	2741	15,53
Московский государственный университет имени М. В. Ломоносова, НИИ физико-химической биологии им. А. Н. Белозерского	2304	18,88
Московский государственный университет имени М. В. Ломоносова, биологический факультет	1985	13,35
Институт молекулярной биологии им. В. А. Энгельгардта РАН	1780	14,11
Московский государственный университет имени М. В. Ломоносова, химический факультет	1375	14,87
Санкт-Петербургский государственный университет	1142	13,07
Институт эволюционной физиологии и биохимии им. И. М. Сеченова РАН	1108	5,53
Новосибирский национальный исследовательский государственный университет	937	8,99
Институт цитологии РАН	805	18,72
Институт органической химии им. Н. Д. Зелинского РАН	804	13,46
Институт белка РАН	772	17,28
Институт теоретической и экспериментальной биофизики РАН	740	11,68
Казанский (Приволжский) федеральный университет	731	8,11
НИИ биомедицинской химии имени В. Н. Ореховича	707	13,36
Институт цитологии и генетики Сибирского отделения (СО) РАН	696	19,61
Институт биохимии им. А. Н. Баха РАН Федерального исследовательского центра (ФИЦ) «Фундаментальные основы биотехнологии» РАН	688	18,76
Московский физико-технический институт (национальный исследовательский университет) (МФТИ)	679	10,21
Институт химической биологии и фундаментальной медицины СО РАН	671	11,18

и скромной аналитики, при неверно введенных данных будут получены неверные результаты, даже если сам по себе алгоритм правилен.

Большое количество неверных интерпретаций данных возникает из-за того, что, например, в РИНЦ потоки первичной и вторичной информации не разделены явным образом, и поэтому возможны парадоксальные ситуации. Рассмотрим такую ситуацию на примере данных советского ученого лауреата Нобелевской премии Виталия Лазаревича Гинзбурга (21.09.1916 –

08.11.2009). В РИНЦ обнаружались публикации за годы как до рождения В. Л. Гинзбурга, так и после его смерти (рис. 2). Справедливости ради стоит отметить, что «лишние» записи перешли из списков пристатейной библиографии, в которых были допущены ошибки. Хотя в РИНЦ и присутствует граница отсечения публикаций из пристатейной библиографии, далеко не все пользователи обращают на нее внимание. Получается, что вторичная и первичная информации слиты, чего быть не должно.

№	Публикация	Цит.
1	Ginzburg V.L., Tsytoovich V.N. Physics Reports, 1919, С. 49.	1
2	КВАНТОВАЯ ТЕОРИЯ СВЕТОВОГО ИЗЛУЧЕНИЯ ЭЛЕКТРОНА, РАВНОМЕРНО ДВИЖУЩЕГОСЯ В СРЕДЕ Гинзбург В.Л. Журнал Русского физико-химического общества, 1904, Т. 10, С. 589.	1
3	SOME REMARKS ON PHASE TRANSITIONS OF THE SECOND KIND AND THE MICROSCOPIC THEORY OF FERROELECTRIC MATERIALS Ginzburg V.L. Physics of the Solid State, 1824, Т. 2, С. 637.	1

Рис. 2. Пример опечаток из источника по РИНЦ. Публикации В. Л. Гинзбурга (1916–2009): публикации до рождения

Fig. 2. The example of misprints of publications in RISC. Publications by V. L. Ginsburg (1916–2009). Mistakes: publications before the birth

На рисунке 2 видно, что публикации В. Л. Гинзбурга цитировались еще до (!) его рождения. Подобные неточности вызывают вопросы и справедливую критику в адрес информационных систем.

Помимо ошибок в описании в самой системе, мы сталкиваемся с проблемой поисковой компетенции пользователя. Учитывая, что далеко не все пользователи понимают различия между первичной и вторичной информацией, а также могут не обратить внимания на необходимость выбора и контроля нужных полей (отмечаются «галочками»), возникают ситуации, которые могут вызывать удивление у неопытных исследователей. Прямой поиск по автору (Гинзбург В. Л.) в РИНЦ дает более 800 публикаций, но, если убрать «галочку» с поля «учитывать публикации, извлеченные из списков цитированной литературы», их число снизится до 177, тогда как WoS CC показывает не более 140 публикаций данного автора. Таким образом, специалисты, занимающиеся библиометрическими оценками, должны владеть всеми тонкостями используемых инструментов и понимать предмет поиска и оценки.

Ошибки и неточности есть во всех базах цитирования, но это больше относится ко вторичной информации. Тем не менее до недавнего времени в Scopus наблюдались случаи, когда наряду с информацией об уже вышедшей публикации присутствовали записи о ней же, но с разными статусами («в печати» и «финальная»), поэтому число выявленных публикаций могло вырасти за счет дублированных документов.

2. Кто несет ответственность за качество ввода данных

Ответственность за качество ввода данных должны нести владельцы и администраторы ресурсов. Попытки переложить ответственность за это на авторов под предлогом борьбы за качество не решают проблемы, так как никто не может обязать ученых отслеживать информацию о своих публикациях во всех БД, искать неточности, владеть необходимыми компетенциями для корректировки данных, отводить этому процессу немалую часть своего рабочего времени. К тому же нет гарантии корректности данных, представленных учеными. Несколько раз мы сталкивались с ситуациями, когда авторы сознательно привязывали к своему профилю публикации коллег-однофамильцев, повышая тем самым библиометрические показатели. Кто должен следить за этим? Процесс проверки, дополнения, исправления, осуществляемый сторонними и тем более заинтересованными лицами, будет бесконечным.

Заключение

Библиометрия превратилась в важнейший инструмент оценки научной и образовательной сфер, став порой основным фактором для принятия решений в управленческих структурах и в определении научной политики. Корректность библиометрических показателей в значительной степени зависит от качества библиографического описания, что может быть обеспечено только квалифицированными специалистами, владеющими всеми необходимыми компетенциями по работе с БД и умеющими грамотно интерпретировать полученные результаты.

Список источников / References

- Арчаков А. И., Карпова Е. А., Пономаренко Е. А. Международные критерии эффективности научно-исследовательской деятельности коллективов и отдельных ученых в области биологии и медицины // Вестник Российской академии медицинских наук. 2013. № 5. С. 4–9 [Archakov AI, Karpova EA and Ponomarenko EA (2013) International criteria for the effectiveness of research activities of teams and individual scientists in the field of biology and medicine. *Vestnik Rossiiskoi akademii nauk* 5: 4–9. (In Russ.)].
- Березкина Н. Ю., Хренова Г. С. Использование баз данных «Web of Science» для оценки результатов научной деятельности в республике Беларусь // Образовательные технологии и общество. 2010. Т. 13, № 3. С. 311–316 [Berezkina NYu and Khrenova GS (2010) Using the “Web of Science” databases to assess the results of scientific activities in the Republic of Belarus. *Obrazovatel'nye tekhnologii i obshchestvo* 13 (3): 311–316. (In Russ.)].
- Гиляревский Р. С. Публикационная активность как оценка научных достижений // Научно-техническая информация. Серия 1. 2014. № 8. С. 1–9 [Gilyarevsky RS (2014) Publication activity as an assessment of scientific achievements. *Nauchno-tekhnicheskaya informatsiya. Seriya 1* 8: 1–9. (In Russ.)].
- Гуреев В., Мазов Н., Каленов Н. База данных RSCI: некоторые оценки наполнения и возможностей использования // Информация и инновации. 2017. Спец. вып. С. 46–51 [Gureev V, Mazov N and Kalenov N (2017) RSCI database: some estimates of content and possibilities of use. *Informatsiya i innovatsii* sp. iss: 46–51. (In Russ.)].
- Гуськов А. Е. Реформа российской науки как импульс для развития наукометрических исследований (вступительная статья) // Труды ГПНТБ СО РАН. 2015. Вып. 9. С. 5–13 [Guskov AE (2015) The Russian science reform as an impetus to develop scientometric research (introductory article). *Trudy GPNTB SO RAN* 9: 5–13. (In Russ.)].
- Лаврик О. Л., Глухов В. А. Публикационная активность автора в РИНЦ: количественные данные как основа для качественного анализа // Труды ГПНТБ СО РАН. 2015. Вып. 9. С. 134–146 [Lavrik OL and Glukhov VA (2015) The author's publication activity in the RSCI: quantitative data as a basis for qualitative analysis. *Trudy GPNTB SO RAN* 9: 134–146. (In Russ.)].
- Маркусова В. А. Проверено временем // Поиск. 2014. 8 авг. (№ 31/32). С. 7 [Markusova VA (2014) Time-tested. *Poisk* 31/32: 7. (In Russ.)].
- Маршакова-Шайкевич И. В. Роль библиометрии в оценке исследовательской активности науки // Управление большими системами. 2013. Спец. вып. 44. С. 210–247 [Marshakova-Shaykevich IV (2013) The role of bibliometrics in assessing the research activity of science. *Upravlenie bol'shimi sistemami* 44: 210–247. (In Russ.)]. URL: http://ubs.mtas.ru/archive/search_results_new.php?publication_id=19060 (дата обращения = accessed 18.04.2021).
- Москалева О. В. Можно ли оценивать труд ученых по библиометрическим показателям? // Управление большими системами. 2013. Спец. вып. 44. С. 308–331 [Moskaleva OV (2013) Is it possible to evaluate the work of scientists by bibliometric indicators? *Upravlenie bol'shimi sistemami* 44: 308–331. (In Russ.)]. URL: http://ubs.mtas.ru/archive/search_results_new.php?publication_id=19065 (дата обращения = accessed 18.04.2021).
- Мохначева Ю. В., Харибина Т. Н. Обеспечение ученых библиометрической информацией в Центральной библиотеке Пушкинского научного центра РАН (отдел БЕН РАН) // Новые технологии в информационно-библиотечном обеспечении научных исследований. Екатеринбург, 2010. С. 218–225 [Mokhnacheva YuV and Kharybina TN (2010) Providing scientists with bibliometric information in the Central Library of the Pushchino Scientific Center of the Russian Academy of Sciences (LES RAS department). *Novye tekhnologii v informatsionno-bibliotechnom obespechenii nauchnykh issledovaniy*. Ekaterinburg, pp. 218–225. (In Russ.)].
- Мохначева Ю. В., Цветкова В. А. Возможности и особенности «Web of Science Core Collection», «Scopus» и РИНЦ для анализа публикационных потоков научных организаций // Информационное обеспечение науки: новые технологии. Москва, 2017. С. 50–64 [Mokhnacheva YuV, Tsvetkova VA (2017) Opportunities and features of “Web of Science Core Collection”, “Scopus” and RSCI for the analysis of publication flows of scientific organizations. *Informatsionnoe obespechenie nauki: novye tekhnologii*. Moscow, pp. 50–64. (In Russ.)].
- Цветкова В. А. Система цитирования: где зло, где благо // Научные и технические библиотеки. 2015. № 1. С. 18–22 [Tsvetkova VA (2015) Citation system: where is evil, where is good. *Nauchnye i tekhnicheskie biblioteki* 1: 18–22. (In Russ.)].
- Цветкова В. А., Мохначева Ю. В., Калашникова Г. В. Парадоксы библиометрических инструментов // Научные и технические библиотеки. 2018. № 8. С. 3–19 [Tsvetkova VA, Mokhnacheva YuV and Kalashnikova GV (2018) Paradoxes of bibliometric tools. *Nauchnye i tekhnicheskie biblioteki* 8: 3–19. (In Russ.)].
- Bernal JD (1939) The social function of science. London: Routledge.